TNO Institute for Perception

TNO Defence Research

AD-A273 761

Kampweg 5
P.O. Box 23
3769 ZG Soesterberg
The Netherlands

Fax +31 3463 5 39 77
Telephone +31 3463 5 62 11

TD 93-1361

1

TNO-report IZF 1993 B-9

ENDOGENOUS AND EXOGENOUS CONTROL OF VISUAL SELECTION: A REVIEW OF THE LITERATURE

J. Theeuwes

Number of pages: 21

Netherlands organization for applied scientific research

TNO Defence Research consists of:
the TNO Physics and Electronics Laboratory, the TNO Prins Maurits Laboratory and the TNO Institute for Perception.

Best Available Copy

CONTENTS Page

Report No.: IZF 1993 B-9

Title: Endogenous and exogenous control of visual selection: a review of the literature

Author: Dr.Ing. J. Theeuwes

Institute: TNO Institute for Perception
Group: Skilled Behavior

Date: July 1993

DO Assignment No.: B92-28

No. in Program of Work: 788.2

SUMMARY

Among the most fundamental issues of visual attention research is the extent to which visual selection is controlled by properties of the stimulus or by the intentions, goals and beliefs of the observer. Before selective attention operates, preattentive processes perform some basic analyses segmenting the visual field into functional perceptual units. The crucial question is whether the allocation of attention to these perceptual units is under the endogenous control of the observer (intentions, goals, beliefs) or under the exogenous control of stimulation. This report discusses evidence regarding the endogenous and exogenous control of attention in tasks in which subjects search for a particular "basic" feature (e.g., search for a unique color, shape, brightness). The present review suggests that selectivity in these type of search tasks is dependent on the relative saliency of the stimulus attributes. It is concluded that the visual system automatically calculates differences in basic features (e.g., difference in shape, color, brightness) and that visual information occupying the position of the highest saliency across stimulus dimensions is exogenously passed on to the "central representation" that is responsible for further stimulus analysis. Alternative explanations of the present findings and tentative speculations resulting from the present approach are discussed.

Rap.nr. IZF 1993 B-9 Instituut voor Zintuigfysiologie TNO
Soesterberg

Interne en externe sturing van visuele selectie: een overzicht van de literatuur

J. Theeuwes

SAMENVATTING

Één van de belangrijkste vragen in het visuele aandachtsonderzoek is de mate waarin visuele selectie bepaald wordt door de eigenschappen van de stimuli aanwezig in het visuele veld of door de intenties van de waarnemer. In het algemeen wordt verondersteld dat pre-attentieve processen, het visuele veld opdelen in functionele perceptuele eenheden. De cruciale vraag is of het richten van aandacht naar deze perceptuele units onder controle staat van de waarnemer of gecontroleerd wordt door stimulatie uit de omgeving. In dit rapport wordt gekeken naar deze interne en externe sturing van aandacht wanneer proefpersonen dienen te zoeken naar unieke zgn. "basic features" zoals kleur, vorm, helderheid, etc. Uit dit overzicht blijkt dat in dit soort zoektaken selectiviteit bepaald wordt door de relatieve discrimineerbaarheid van de stimulus attributen. Het visuele systeem berekent automatisch de verschillen in de basic features en het object op de locatie met de hoogste opvallendheid wordt automatisch geselecteerd. Alternatieve verklaringen en nieuwe speculaties worden in dit rapport besproken.

1 INTRODUCTION

The ability to detect and/or recognize objects in the visual environment plays an essential adaptive role for human behavior, in particular for acting in a goal-directed manner. It is commonly known, however, that, at any one time, one can process only a small amount of information present in the visual field. This limitation stresses the importance of selection: at any one time, it is important to select those objects needed to guide current behavior. The limitation in processing implies that, at some stage (or stages) in visual information processing, some objects are selected for further processing while others are excluded. This process of selecting part of simultaneous sources of information, either by enhancing the processing of some objects and/or by suppressing information of others, is traditionally referred to as *"selective attention"* (Johnston & Dark, 1986). Theories of human selective attention are concerned with how people select information to provide the basis for responding and with how information, irrelevant to that response, is dealt with.

Most current accounts of selective attention theories suggest that selection is controlled in two distinct ways. When an observer intentionally selects from the visual field only those objects which are required to perform the task at hand, selection is thought to occur in a *goal-directed*, voluntary manner. When specific properties present in the visual field capture attention independently of the observer's goals and beliefs, selection is thought to occur in *stimulus-driven*, involuntary manner. These two mechanisms of selection have been referred to as *endogenous and exogenous* control, respectively (Posner, 1980; Folk, Remington & Johnston, 1992). Visual selection may be controlled by either one of these systems or a combination of them (Yantis, in press a).

Visual selection can only be involved when simultaneous sources of information compete for *selection*. In other words, selection can only occur when an observer has to select one object among different other objects. The flow of information runs from distinct objects present in the visual field to a single response. Selection determines which object (or objects) is processed first, second, third, etc. It is generally assumed that before selective attention operates, preattentive processes perform some basic analysis segmenting the visual field into functional perceptual units. Directing attention to one of these units implies that such a unit has been *selected* for further more sophisticated processing (Broadbent, 1958, 1982).

The dichotomy between an early preattentive process that segments the visual field into basic units followed by a second attentive stage which processes information in more detail is recognized by most current accounts of human vision [e.g., Feature Integration Theory (Treisman & Gelade, 1980; Treisman, 1988; Treisman & Sato, 1990) Julesz's "texton" theory (Bergen & Julesz, 1983; Julesz, 1971), Cave and Wolfe's "guided search" model (Cave & Wolfe, 1990; Wolfe, Cave & Franzel, 1989), Hoffman's two-stage model (1978, 1979), and

Duncan and Humphreys' similarity theory (Duncan & Humphreys, 1989)]. Preattentive segmentation is thought to occur without capacity limitations in parallel across the visual field, whereas the attentive processing requires the allocation of attentional resources to a location in visual space. The latter processing system has a limited capacity and processes information serially.

The dichotomy between these two processes typically shows up in visual search tasks, in which an observer is asked to determine whether a target stimulus is present among a variable number of distractor stimuli. The total number of stimuli present in the display is usually referred to as the *display set size*. In tasks in which an observer has to detect a target defined by a primitive feature such as color, shape, size and brightness, there is hardly a set-size effect (e.g., Egeth, Jonides & Wall, 1972). Typically, search functions with slopes which are less than 5- or 6- ms per item are considered to reflect parallel search (Treisman & Souther, 1985). Such a "pop-out effect" is used as a diagnostic that the information that defines the target is available at the preattentive parallel level (Treisman & Gormican, 1988). For example, a red object embedded in an array of green distractors will pop-out, that is, the time to detect the red object is independent of the number of green objects. In terms of the framework described above, it is assumed that the preattentive parallel stage segments the visual field in a single red and a group of green items. Although the presence of the red item is coded at the preattentive parallel stage, it is assumed that the target item should enter the second attentive stage of processing before a response can be given (e.g., Johnston & Pashler, 1990; Theeuwes, 1993a; Tsal & Lavie, 1993; yet see Folk & Egeth, 1989). In other words, following preattentive segmentation, spatial attention is shifted to the location of the red item, implying that the red item enters the second stage of attentive processing.

Search functions reflecting parallel search can be contrasted with search functions showing a linear increase in search time with the number of non-target items in the display. Usually the slope of target absent trials is twice as steep as the slope of target present trials suggesting spatially serial, self-terminating search (Treisman & Gelade, 1980; Quinlan & Humphreys, 1987). This pattern of results typically shows up in case the target is defined by *conjunctions* of elementary features. For example, search for a vertical, red line segment between tilted red line segments and vertical green line segments will give serial search functions. Because display elements can only be classified as targets and non-targets by means of the second—limited capacity—stage of attentive processing, serial scanning through the display is necessary giving a large effect of the number of nontargets on search times. It should be noted that also in cases of conjunction search, it is likely that some preattentive segmentation at a featural level will take place parsing the visual field in different groups of items. In the example above, it is likely that two different segmentations occur: one in the color dimension (green versus red) and one in the orientation dimension (tilted versus vertical).

Recent theories of visual search recognize the initial segmentation and assume that this segmentation might "guide" search for conjunction targets (e.g., Egeth, Virzi & Garbart, 1984; Wolfe et al., 1989; Treisman & Sato, 1990; Zohary & Hochstein, 1989). These notions are supported by empirical evidence showing that there is not always a clear difference between parallel and serial search functions (e.g., Duncan & Humphreys, 1989; Nakayama & Silverman, 1986; Wolfe et al., 1989).

In the descriptions above, the implicit assumption is made that attending to a stimulus location has a special status (see also, Van der Heijden, 1992). Attentive processing is equal to directing spatial attention to a location in the visual field. Thus, serial attentive processing is equivalent to directing the "spotlight of attention" (Posner, 1980) serially to locations in space. Recent evidence confirms the notion that selection is always based on spatial location. Tsal & Lavie (1988, 1993) show that attending to any aspect of a stimulus (attend to color, attend to shape) automatically entails directing attention to the stimulus' location. This result suggests that directing attention to a location in space is not merely necessary to conjoin features (as for example advocated by Treisman & Gelade, 1980), but it is a mandatory process occurring both during feature or conjunction search regardless of the particular property to which the observer tries to attend.

Recently a considerable debate has erupted regarding the extent to which *selection* in visual search is controlled exogenously or endogenously (e.g., Bacon & Egeth, 1993; Duncan & Humphreys, 1989; Folk, Remington & Johnston, 1992, in press; Theeuwes, 1991a, 1992, 1993a, 1993b; Wolfe et al., 1989; Yantis, in press a, in press b; Yantis & Jonides, 1984). As outlined above *selection* means that a particular item enters the second stage of attentive processing. The crucial question is whether it is possible to exert top-down control over the preattentive parallel stage of processing so that only information required to perform the task at hand enters the second stage of processing or whether the physical properties of the stimuli present in the visual field dictate what will and will not enter the second stage of processing. In other words: is selection in visual search the result of endogenous control of the observer (intentions, goals, beliefs) or is it the result of the exogenous control of stimulation?

In this report, I critically examine evidence for exogenous and endogenous control of attention. Relevant empirical evidence will be discussed and when necessary reinterpreted in the context of the issue. Finally, I will propose a parsimonious account for the findings on exogenous and endogenous control of selection and will discuss some tentative speculations.

2 EXOGENOUS AND ENDOGENOUS CONTROL OF ATTENTION

Exogenous control refers to the condition in which selection is determined by the attributes of the stimulus and not by the observer's goals or intentions. For example, when an observer is confronted with a display with one red and several green items, it is probable that the red item is automatically selected, that is, the red item automatically pops-out and enters the second stage of attentive processing. When such a result is found, one might conclude that the selection of the red item was under the control of stimulation, and occurred automatically, similar to the belief that when one is looking around, conspicuous objects "demand to be looked at" (e.g., Engel, 1977). Yet, in the example above it is unclear what attentional set the observer adopted, that is, it is possible that the observer deliberately looked for red items. Jonides and Yantis (1988) and Theeuwes (1990) investigated this issue. When an observer was searching for a target which could *not* be detected preattentively (e.g., Jonides & Yantis, 1988: looking for the letter E between a varying number of other letters), the presence of an irrelevant featural singleton in color, brightness or shape did not affect search behavior. The results showed that the featural singleton was ignored and search time increased linearly with the number of elements in the display. Theeuwes (1990, 1993a, 1993b) suggested that these featural singletons did not affect search behavior because the target (the letter E) could only be detected among the nontarget letters by means of the second stage of attentive processing stage, that is, serial scanning through the display was necessary in order to detect the target. Because the observer knew that the target could not be detected by means of pre-attentive processing (e.g., a letter E does not pop-out among other letters), it was claimed that observers adopted a strategy that allowed them to immediately start processing the display at the attentive level. Since attentive processing is equal to the direction of spatial attention to a location in the visual field, it is hypothesized that observers might have focussed in on a particular location and serially checked the locations for the target element. As a consequence of this particular attentional "serial search" strategy, it is speculated the preattentive parallel segmentation process was bypassed. Therefore, the irrelevant featural singleton could not have had an effect on performance because the singleton was not segmented from the other elements. These findings suggest that when serial attentional scrutiny is required, the adopted top-down strategy can override stimulus-driven capture by a (static) featural singleton.

Yantis and Jonides (1984) however showed that under the same circumstances when subjects have to serially search through a display, irrelevant dynamic discontinuities (e.g., abrupt onsets) are always selected first. In Yantis and Jonides (1984), on each trial, one letter had abrupt onset. When the abrupt-onset letter was the target, search time became independent of the number of distractors, suggesting that the element with abrupt onset always entered the second stage of attentive processing first. In addition, Theeuwes (1990, Exp. 3) showed that an irrelevant element with abrupt change (e.g., an element was changed from a square to a circle) tends to be selected first on about 25% of the trials.

These results indicate that dynamic discontinuities are special in the sense that they occasionally capture attention even when subjects have the intention to search serially. A possible neural mechanism for the special status of dynamic discontinuities is provided by Breitmeyer and Ganz (1976). They claim that transient channels in the visual system which selectively respond to onsets and offsets transmit their signal rapidly to the brain.

The idea that observers can use an attentional set that allows them to process information at the attentive level only (like the "serial search" strategy) is confirmed by experiments that show that even dynamic discontinuities do not capture attention when subjects are in a spatially focussed state (Yantis & Jonides, 1990; Theeuwes, 1991b). When subjects endogenously focus their attention to a cued location in visual space, irrelevant abrupt onsets and offsets presented elsewhere in the visual field do not capture attention anymore. It is hypothesized that the information present at the location to which subjects focus their attention enters directly the second stage of attentive processing. Again, preattentive processing which might have signalled the abrupt onset or offset is passed by.

The account above suggests a great deal of endogenous control over visual selection. Yet, the analysis indicates that the way endogenous control is obtained is rather limited. Subjects can only exert control over visual selection through the second stage of attentive processing. By varying the size of the attended area (e.g., spotlight or zoom-lens metaphors of attention; Eriksen & Yeh, 1985; Humphreys, 1981; Posner, 1980), the area in which preattentive segmentation can occur varies as well. Endogenously directing attention to a location operates as a top-down spatial filter: information outside the attended area does not enter into the system, that is it does not enter the second stage of attentive processing. The present account is in line with the claim that location is special and that selection is always based on a spatial location (Tsal & Lavie, 1993). The claim is made here that the control of attention can be completely top-down in a sequential focussed search of single items or groups of items.

The question remains whether it is possible to have goal-directed selection for non-spatial attributes (e.g., color, shape, brightness) in cases in which attention is not focussed. If preattentive parallel search occurs (e.g., searching for a singleton item which can be detected preattentively), is it possible to select only items which are relevant for the task? The question is simple: if an observer is looking for a circle between several squares, can he/she endogenously alter the system so that only the circle enters the second stage of processing (i.e., that only the circle is *selected*). Theeuwes (1991b, 1992) examined this question by conducting a visual search task in which subjects had a clear attentional set to select a singleton (for example, the target is a green circle and the nontargets are green squares). Subjects viewed multi-element displays (5, 7, 9 items) and had to respond to the orientation of a line segment (horizontal or vertical) that always appeared inside a green circle (see Fig. 1).

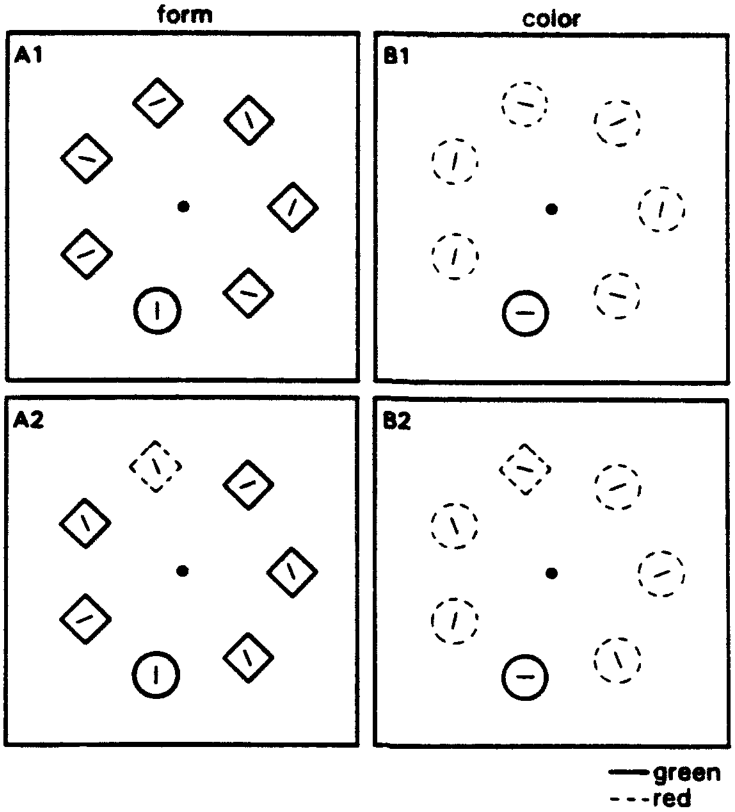

Fig. 1 Subjects cannot ignore an irrelevant singleton when searching for a singleton. Top: Subjects had to report the orientation of the single nonoblique line segment. In form singleton condition, the target line segment was always inside the green circle surrounded by green squares (left panels). In the color singleton condition, the target line segment was always inside the green circle surrounded by red circles (right panels). In the top panels (A1 and B1), there is no distractor. In the bottom panels (A2 and B2) there is a distractor item.

Nontarget line segments appeared inside either green squares (form condition: the target was a form singleton, see Fig. 1 panels A1 and A2) or red circles (color condition: the target was a color singleton, see Fig. 1, panels B1 and B2). In each of these conditions, a known, irrelevant singleton distractor in the other dimension than the relevant one was present on half of the trials. In the form condition half of the trials did not contain a distractor (panel A1), and the other half contained a red square in addition to the green target circle and the nontarget green squares (panel A2). In the color condition half of the trials did not contain a distractor (panel B1), and the other half contained a red square in addition to the green target circle and the nontarget red squares. The presence of an irrelevant color singleton (the red square) when one is looking for a shape singleton (a circle among squares) significantly elevated reaction time (Fig. 2, panel A).

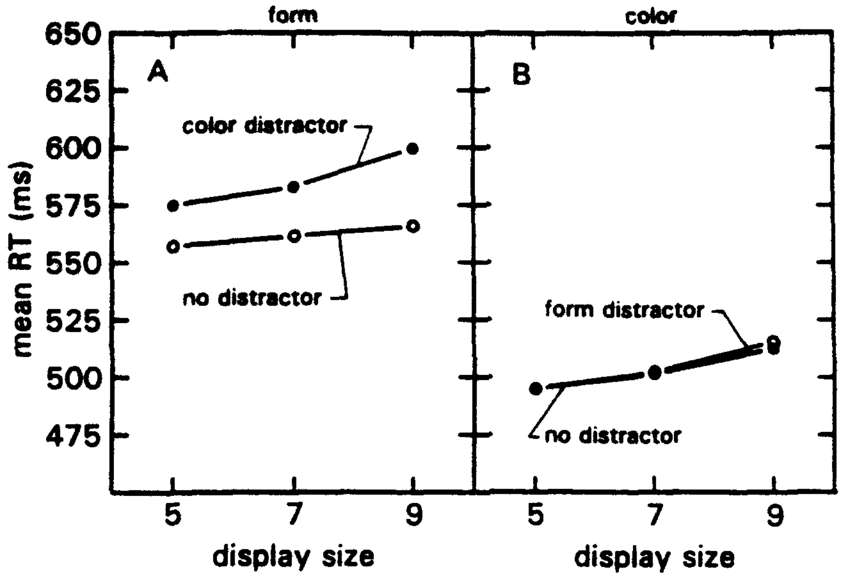

Fig. 2 Reaction time as a function of display size for search with or without a distractor for the form (Panel A) and color (Panel B) conditions. From Theeuwes (1992, Exp. 1A).

The results show that when one is searching for a known singleton (in this case, a target green circle), a salient singleton known to be irrelevant (in this case, a red square), will capture attention. The absence of an effect of display size (flat search functions) is important because it indicates that subjects did not use a sequential focussed search mode, which, as discussed earlier, diminishes distracting effects of events falling outside the attentional beam (e.g., Theeuwes, 1991a). The results indicate that even though the observer knows that the singleton is irrelevant he/she cannot help that this singleton enters the second stage of attentive processing first. After entering the attentive processing stage, the distractor will be discarded quickly because subjects know that they are looking for a green circle and not for a red square. Attention will be disengaged and switched to the next singleton which in this case is the target. This is clearly a top-down effect; yet, note that this effect operates on a *selected* item, i.e. it is an effect that operates on the second attentive stage of processing.

Note that the capturing of attention of the irrelevant singleton is a robust effect. Theeuwes (1992, exp. 1A) trained subjects for almost 2000 trials. Even after this extended and consistent practice subjects lacked the ability to simply ignore the known-to-be-irrelevant color singleton.

Panel B of Fig. 2 however shows that not every singleton captures attention: the presence of an irrelevant shape singleton (the green square) did not affect search for the color singleton (a green between red items). Rather than assuming that this successful selection of the target singleton is due to a top-down altering of

the system, the data hint towards an alternative explanation. The "no-distractor" conditions shown in Fig. 2 reveal that finding a green circle between red circles (Panel B) is about 60 ms faster than finding the same stimulus surrounded by green squares (Panel A). This implies that the color singleton is more salient than the form singleton. If the most salient singleton captures attention first, then the asymmetric selectivity depicted in Fig. 2 (Panel A: a color singleton interferes with search for a form singleton; Panel B: a form singleton does not interfere with search for a color singleton) can be explained without assuming any top-down control.

Theeuwes (1991b, Exp. 3 and 1992, Exp. 2) tested this notion in an analogous visual search experiment in which the color singleton was less salient (a yellowish green singleton between yellowish red nontargets) than the form singleton. If attention is captured first by the most salient singleton irrespective of whether it is the target or the distractor then one expects to find a reversed selectivity. Fig. 3 shows that this is indeed the case: when searching for a form singleton the relatively less salient color singleton does not interfere (Panel A); on the other hand, when searching for a less salient color singleton, the relatively more salient form singleton does capture attention first and thus elevate response times (panel B).

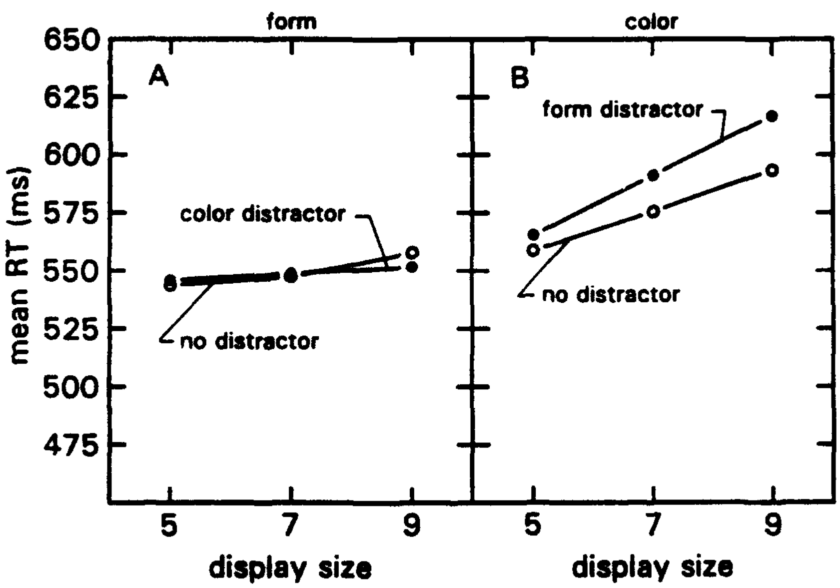

Fig. 3 Reaction time as a function of display size for search with or without a distractor for the form (Panel A) and color (Panel B) conditions. From Theeuwes (1992, Exp. 2).

Note that a less salient color singleton (Fig. 3; Panel B; no distractor condition) still pops-out, yet the time it takes before it pops-out is about 70 ms is longer than when a salient color singleton is used (Fig. 2; Panel B).

Theeuwes' (1992) findings on color and form have recently been replicated by Bacon and Egeth (1993, Exp. 1). Pashler (1988, Exp. 6) using a related paradigm also showed large interference effects for the detection of a target defined by form when a single irrelevant item with a unique color was present. Recently, Theeuwes (1993b) showed that the observed interference effects are not limited to between-dimensions static discontinuity like form and color (Theeuwes, 1991b, Exp. 2 and 3; 1992, Exp. 1 and 2) or brightness and color (Theeuwes, 1991b, Exp. 1) but that similar interferences are found between static discontinuity (color) and dynamic discontinuities (abrupt onsets). Thus, using a similar paradigm, depending on the relative saliency an irrelevant abrupt onset singleton interfered with search for a color singleton and vice versa.

These findings have led to the conjecture that there is no top-down control at the level of preattentive processing. When using the preattentive parallel search mode, the extent to which singletons capture attention is determined by the relative saliency of the singletons present in the visual field. Irrespective of what subjects are looking for (i.e., irrespective of any top-down control), spatial attention is automatically and involuntary captured by the most salient singleton. The shift of spatial attention to the location of the singleton, implies that the singleton is selected for further processing. If this singleton is the target, a response is given. If it is not the target, attention is automatically switched to the next salient singleton.

According to the present notion, the preattentive process simply calculates differences in features within dimensions. This results in a pattern of activations at different locations. For example, at the location of the red singleton a large "difference" signal arises because the singleton differs from all other nontargets in color. At the location of the circle singleton, a large "difference" signal arises because the circle differs from all other elements in shape. Focal attention is automatically and unintentionally shifted to locations in the display containing large local feature differences, regardless of the dimension in which this feature difference occurs. The source of the pre-attentively calculated difference signal (whether it is caused by a color singleton or a form singleton) can only be recognized after attention is moved to the location of the difference signal. In other words, the subject only knows whether the singleton was the target after selecting the location having the large difference signal. In this view, the salience of the singleton, and not its identity, its color, its shape, its brightness, etc., will determine which element captures attention. Obviously, given this model, selection operates irrespective of the task demands. The automatic shifts of attention are considered to be the result of relatively inflexible, "hardwired" mechanisms which are triggered by the presence of these difference signal interrupts. In line with for example Sagi and Julesz (1985) and Ullman (1984) it is assumed that the parallel process can only perform a *local-mismatch* detection followed by a serial stage in which the most mismatching areas are selected for further analysis.

3 FOCUSSING OF ATTENTION AS A FILTERING DEVICE

The analysis above leads to the following conclusions regarding top-down and bottom-up control: (1) when parallel preattentive search is used to detect the target (e.g., the target of search is a featural singleton) selection is completely determined by the saliency of the singletons, (2) when serial attentive search is used (e.g., the target is not a preattentively available singleton), selection is primarily determined by the goals and intentions of the observer (with the exception of abrupt onsets and offsets which occasionally capture attention). Top-down selection can only be based on spatial location, and not on non-spatial attributes like color, shape, brightness etc.

It is speculated that the two modes of selection described above can also work together. For example, subjects may choose to search a display partially serially, in which the size of the attended area is endogenously varied. When the size of the attentional spotlight is reduced, preattentive segmentation within groups of items may take place and within groups of items targets may be detected in parallel. When subjects know they have to search for a salient singleton (as in Theeuwes' experiment described above), the attentional window will be set to cover the whole visual field. As a consequence, preattentive segmentation within that attended area will take place and top-down selectivity within that area is lost, i.e., the most salient item will be selected first. If subjects look for items that do not stand out from the environment they may adopt a smaller attentional window. For example, when searching for a conjunction target, a spatial window that covers for example three groups of items within which a target may pop-out, will give relatively fast search times. Ultimately, when the difference between target and distractors is so small that attentive processing is required to detect this difference (a low signal-to-noise ratio between target and distractor) the attentional window may be so small that it covers individual items (e.g, the sequential focussed search mode). The endogenous controllable variable-size attentional window acts like a early-spatial filter, restricting (preattentive) processing of items within the attended region and blocking out information from all other parts in the visual field (e.g., Yantis & Johnson, 1990; Yantis & Jonides; 1990, Theeuwes, 1991a). In this way top-down control over visual selection is accomplished by a variable-size spatial window (see also, Humphreys & Müller, 1992; Treisman & Gormican, 1988).

4 FURTHER SPECULATIONS

4.1 The strength of the difference signal

The present approach assumes that within the variable-size spatial window, differences in feature dimensions (e.g., difference with the color dimensions, shape dimension, etc.) are calculated automatically. This results in a pattern of

activations displaying difference signals which indicate how different each item is from each of the other items within a particular feature dimension. It is assumed that the calculations between dimensions are independent. Therefore, the strength of the difference signal does not sum up between dimensions, at least not at the preattentive level. Thus, a target that differs from nontarget items both in color and in shape should not produce a larger difference signal than an item that only differs from the other items in color.

4.2 **Topographic information is preserved**

The original feature integration theory assumes that features are represented independent of their locations. Under circumstances of attention overload, these free-floating features may be miscombined into illusory conjunctions, objects consisting of features from different locations (e.g., Treisman & Schmidt, 1982). The present approach which claims that preattentive segmentation only occurs within a variable size window has to assume that the topographic representation of features is preserved (see also, e.g., Green, 1991). As a consequence of a topographic representation, it is likely that the calculation of the difference signal depends on the spatial distance between a singleton and the display elements. Thus, display elements directly neighboring a singleton will contribute more to the difference signal than elements further away. As recognized by Green (1991) this implies that in search tasks it should be possible to find search times which *decrease* with increasing number of display elements because close proximity between the items will make comparisons easier. In fact, with displays as described above (see Fig. 1), Theeuwes (1991b) found small negative search functions when searching for a uniquely colored item (Exp. 1: -2.5 ms/item, and Exp. 2: -2.6 ms/item).

4.3 **Tagging of items having a particular saliency**

Sagi and Julesz (1985) showed that detecting and counting 1 to 4 targets that differ in orientation can be done in parallel by preattentive processing, while identifying the source of the local discontinuity required serial focal attention. These findings suggest that the local differences which are detected by the preattentive process are used to drive the attentional focus from one location having a high local difference signal to the next. At any time one needs to know where one is and where one is going (e.g., Trick & Pylyshyn, 1993). Also, in Theeuwes' (1992) experiments described above, in which the more salient color singleton is selected first and the form singleton is selected second, information regarding the locations of local differences should be preserved. It is assumed that the local activations caused by the differences among the elements are preserved. As for example suggested by Yantis and Jones (1991; Yantis & Johnson, 1990), a priority map representing the current priority tag strength of each element in the scene, might drive focal attention through a scene (see also,

Ullman, 1984). As suggested by Yantis and Jones (1991) the strength of these tags may decay over time. After directing attention to one of the tagged locations, information regarding the item at that location becomes available (e.g. its identity, color, brightness, etc.), and the priority tag of that element will be purged. This purging ensures that this element is not selected again. Note that after selecting an item having a particular priority tag, all elements having the same priority tag might be discarded as well.

5 CONCLUDING COMMENTS

The data-driven selection account as described above is not in accordance with various recent models of visual search which assume that visual selection is the result of top-down and bottom-up effects (e.g., Hoffman, 1978; Treisman & Sato, 1990, Wolfe et al., 1989). Generally, these models took the original feature integration theory (FIT, Treisman & Gelade, 1980) as a starting point and added a new turn: the output of the preattentive stage can guide the attentive serial search. For example, in Wolfe et al.'s (1989) view, *during preattentive parallel search*, knowing that one is looking for a green circle, is supposed to enhance the activity of green and circular elements. Because the activity of likely targets is heightened during preattentive processing, attentive serial search will be directed to likely target candidates only. These "guided search" notions assume top-down effects on preattentive parallel search. It should be noted that top-down control is assumed in order to account for relatively flat search functions when searching for targets defined by conjunctions of features. The notion that relatively flat search functions for conjunction targets necessarily represent top-down guided search can be questioned. It is likely that preattentively the display is parsed in groups of items. This parsing of the visual field is assumed to take place without any top-down control. If search is serial between and parallel within these groups then the increase in RT with increasing numbers of items (e.g., 1 to 12 items) might reflect an increase in scanning one to three preattentively parsed groups of items. Obviously, search functions will be rather flat (e.g., going from one to three groups); yet, one does not need to conclude that top-down activation guided attention to those elements that are most similar to the target.

Along similar lines, Bacon and Egeth (1993) showed that when subjects are looking for a particular feature which is not unique within a display, subjects employed a so called "feature search mode". When employing this search mode, search is partially serial through the display (small positive search functions) thereby blocking out the distracting effects of irrelevant singletons (see section 3: focussing of attention as a filtering device). When subjects searched for a specific target feature which was unique in the display (e.g., a green circle among green diamonds as in Theeuwes, 1992), a singleton known to be irrelevant (e.g., a red diamond) captured attention. On the basis of these findings Bacon and Egeth (1993) suggested two different search modes: a "feature search mode" in which

subjects search for a specific shape and a "singleton search mode" in which subjects look for the odd-man-out. Only in this latter search mode, top-down control is not possible: any feature that stands out from the environment attracts attention (see also Wolfe & Cave, 1990, p. 92-93).

A recent study conducted by Folk, Remington and Johnston (1992, in press) challenged the presently advocated data-driven selection account altogether: Folk et al. (1992) claim that selection is never purely stimulus-driven but is always dependent on the internal control settings. In their experiments, subjects had to *ignore* cues immediately prior to the presentation of the target display. It was demonstrated that an onset singleton serving as a cue, does not capture attention when observers adopt an attentional set for color singletons. On the other hand, when observers are set to identify a color singleton, they cannot ignore another color singleton known to be irrelevant (the cue). Folk et al. conclude that all attentional shifts are mediated by "programmable" control settings. Because Folk et al. conclusions are important, Theeuwes (1993b) tried to replicate their findings by means of a more conventional search task similar to the one described earlier (see Fig. 1). Subjects searched multielement displays in which a color singleton and an onset singleton were simultaneously present. When subjects search for a color singleton, on some trials another location contained an irrelevant onset. In addition, when subjects had to search for an onset singleton, on some trials another location contained an irrelevant singleton. The results showed that the Folk et al.'s claim that attentional capture was contingent on internal (top-down) control settings did not hold: in line with earlier findings, Theeuwes (1993b) showed that irrespective of any internal control settings, attention was captured by the most salient event.

There are various reasons why Theeuwes (1991b, 1992, 1993b) has found no evidence of top-down control at the preattentive level whereas others do claim to have obtained such results. First, because the interference effects are relatively small (about 15 to 25 ms), the addition of noise to the display will obscure the interference effect, especially because the conclusion that there is no interference is reached by accepting a null effect. Second, in order to disclose interference effects at the preattentive parallel level, it must be ensured that search is performed in parallel. If search is partially serial (e.g., serial search between and parallel search within clumps of items) as for example with conjunction search then the effect of the distractor will be attenuated. Third, the paradigm used by Theeuwes (1991b, 1992, 1993b) ensured that there is a clear separation between perceptual and response-selection factors. Because subjects responded to the orientation of the target line segment located in the singleton, the stimulus information separating the target from nontargets tells nothing about which of the possible responses to choose. In other words, RT data reflect effects operating at the early stage of *perceptual* processing rather than on processing operations occurring after the item has already been selected (after entering the second stage of processing). For example, knowing the task-relevant stimulus feature might speed up the identification of an item *that has already been*

selected, similar as a prime speeds up processing of a target in a typical priming experiment.

REFERENCES

Bacon, W.F. & Egeth, H.E. (1993). Overriding stimulus-driven attentional capture. Submitted for publication.

Bergen, J.R. & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. Nature 303, 696-698.

Breitmeyer, B.C. & Ganz L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. Psychological Review 83, 1-36.

Broadbent, D.E. (1958). Perception and communication. London: Pergamon.

Broadbent, D.E. (1982). Task combination and the selective intake of information. Acta Psychologica 50, 253-290.

Bundesen, C. (1990). A theory of visual attention. Psychological Review 97, 523-547.

Cave, K.R. & Wolfe, J.M. (1990). Modeling the role of parallel processing in visual search. Cognitive Psychology 22, 225-271.

Duncan, J. & Humphreys, G.W. (1989). Visual search and stimulus similarity. Psychological Review 96, 433-458.

Egeth, H., Jonides, J. & Wall S. (1972). Parallel processing of multielement displays. Cognitive Psychology 3, 674-698.

Egeth, H.E., Virzi, R.A. & Garbart, H. (1984). Searching for conjunctively defined targets. Journal of Experimental Psychology: Human Perception and Performance 10, 32-39.

Engel, F.L. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. Vision Research 17, 95-108.

Eriksen, C.W. & Yeh, Y.Y. (1985). Allocation of attention in the visual field. Journal of Experimental Psychology: Human Perception and Performance 11, 583-597.

Folk, C.L. & Egeth, H. (1989). Does the identification of simple features require serial processing? Journal of Experimental Psychology: Human Perception and Performance 15, 97-110.

Folk, C.L., Remington, R. & Johnston, J.C. (1992). Involuntary covert orienting is contingent on attentional control settings. Journal of Experimental Psychology: Human Perception and Performance 18, 1030-1044.

Folk, C.L., Remington, R. & Johnston, J.C. (in press). Attentional control setting: A reply to Yantis (1993). Journal of Experimental Psychology: Human Perception and Performance.

Green, M. (1991). Visual search, visual streams, and visual architectures. Perception & Psychophysics 50, 388-403.

Hoffman, J.E. (1978). Search through a sequentially presented display. Perception & Psychophysics 23, 1-11.

Hoffman, J.E. (1979). A two-stage model of visual search. Perception & Psychophysics 25, 319-327.

Humphreys, G. W., (1981). On varying the span of visual attention: evidence for two modes of spatial attention. Quarterly Journal of Experimental Psychology 33A, 17-31.

Humphreys, G.W. & Müller, H.J. (1992). Search via recursive rejection (SERR): A connectionist model of visual search. Cognitive Psychology 25, 43-110.

Jonides, J. & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. Perception & Psychophysics 43, 346-354.

Johnston, J.C. & Pashler, H. (1990). Close binding of identity and location in visual feature perception. Journal of Experimental Psychology: Human Perception and Performance 16, 843-856.

Johnston, W.A. & Dark, V.J. (1986). Selective attention. Annual Review of Psychology 37, 4.

Julesz, B. (1971). Foundations of Cyclopean Perception. Chicago: University of Chicago Press.
Nakayama, K. & Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. Nature 320, 264-265.
Pashler, H. (1988). Cross-dimensional interaction and texture segregation. Perception & Psychophysics 43, 307-318.
Posner, M. (1980). Orienting of attention. Quarterly Journal of Experimental Psychology 32, 3-25.
Quinlan, P.T. & Humphreys, G.W. (1987). Visual search for targets defined by combinations of color, shape and size: An examination of the task constraints on feature and conjunction searches. Perception & Psychophysics 41, 455-472.
Sagi, D. & Julesz, B. (1985). Detection versus discrimination of visual orientation. Perception 14, 619-628.
Theeuwes, J. (1990). Perceptual selectivity is task-dependent: evidence from selective search. Acta Psychologica 74, 81-99.
Theeuwes, J. (1991a). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. Perception & Psychophysics 49, 83-90.
Theeuwes, J. (1991b). Cross-dimensional perceptual selectivity. Perception & Psychophysics 50, 184-193.
Theeuwes, J. (1992). Perceptual selectivity for color and form. Perception & Psychophysics 51, 599-606.
Theeuwes, J. (1993a). Visual selective attention: a theoretical analysis. Acta Psychologica 83, 93-154.
Theeuwes, J. (1993b). Stimulus-driven capture and attentional set: selective search for color and visual abrupt onsets. Manuscript submitted for publication.
Treisman, A.M. (1988). Feature and Objects: The fourteenth Bartlett memorial lecture. The Quarterly Journal of Experimental Psychology 40, 201-237.
Treisman, A.M. & Gelade, G. (1980). A feature integration theory of attention. Cognitive Psychology 12, 97-136.
Treisman, A.M. & Gormican, S. (1988). Feature search in early vision: Evidence from search asymmetries. Psychological Review 95, 15-48.
Treisman, A.M., & Sato, S. (1990). Conjunction search revisited. Journal of Experimental Psychology: Human Perception and Performance 16, 459-478.
Treisman, A.M. & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. Cognitive Psychology 14, 107-141.
Treisman, A.M. & Souther, J., (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. Journal of Experimental Psychology: General 114, 285-310.
Trick, L.M. & Pylyshyn, Z.W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. Journal of Experimental Psychology: Human Perception and Performance 19, 331-351.
Tsal, Y. & Lavie, N. (1988). Attending to color and shape: the special role of location in selective visual processing. Perception & Psychophysics 44, 15-21.
Tsal, Y. & Lavie, N. (1993). Location dominance in attending to color and shape. Journal of Experimental Psychology: Human Perception and Performance 19, 131-139.
Van der Heijden, A.H.C. (1992). Selective attention in vision. London: Routledge.
Ullman, S. (1984). Visual routines. Cognition 18, 97-159.
Wolfe, J.M. & Cave, K.R. (1990). Deploying Visual Attention: The guided search model. In: Blake, A. & Troscianko, T.: Ai and the Eye. Chichester: John Wiley & Sons Ltd.

Wolfe, J.M., Cave, K.R. & Franzel, S.L. (1989). Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance 15, 419-433.

Yantis, S. (in press a). Stimulus driven attentional capture and attentional control settings. Journal of Experimental Psychology: Human Perception and Performance.

Yantis, S. (in press b). Stimulus-driven attentional capture. Current Directions in Psychological Science.

Yantis, S. & Johnson, D.N. (1990). Mechanism of attentional priority. Journal of Experimental Psychology: Human Perception and Performance 16, 812-825.

Yantis, S. & Jones, E (1991). Mechanism of attentional selection: Temporally modulated priority tags. Perception & Psychophysics 50, 166-178.

Yantis, S. & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from selective search. Journal of Experimental Psychology: Human Perception and Performance 10, 601-621.

Yantis, S. & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. Journal of Experimental Psychology: Human Perception and Performance 16, 121-134.

Zohary, E. & Hochstein, S. (1989). How serial is serial processing in vision. Perception 18, 191-200.

Soesterberg, July 15, 1993

Dr.Ing. J. Theeuwes

REPORT DOCUMENTATION PAGE

1. DEFENCE REPORT NUMBER (MOD-NL) TD 93-1361	2. RECIPIENT'S ACCESSION NUMBER	3. PERFORMING ORGANIZATION REPORT NUMBER IZF 1993 B-9
4. PROJECT/TASK/WORK UNIT NO. 788.2	5. CONTRACT NUMBER B92-28	6. REPORT DATE 15 July 1993
7. NUMBER OF PAGES 21	8. NUMBER OF REFERENCES 55	9. TYPE OF REPORT AND DATES COVERED Final

10. TITLE AND SUBTITLE

Endogenous and exogenous control of visual selection: a review of the literature

11. AUTHOR(S)

J. Theeuwes

12. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

TNO Institute for Perception
Kampweg 5
3769 DE SOESTERBERG

13. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

TNO Defence Research
Schoemakerstraat 97
2628 VK Delft

14. SUPPLEMENTARY NOTES

15. ABSTRACT (MAXIMUM 200 WORDS, 1044 BYTE)

Among the most fundamental issues of visual attention research is the extent to which visual selection is controlled by properties of the stimulus or by the intentions, goals and beliefs of the observer. Before selective attention operates, preattentive processes perform some basic analyses segmenting the visual field into functional perceptual units. The crucial question is whether the allocation of attention to these perceptual units is under the endogenous control of the observer (intentions, goals, beliefs) or under the exogenous control of stimulation. This report discusses evidence regarding the endogenous and exogenous control of attention in tasks in which subjects search for a particular "basic" feature (e.g., search for a unique color, shape, brightness). The present review suggests that selectivity in these type of search tasks is dependent on the relative saliency of the stimulus attributes. It is concluded that the visual system automatically calculates differences in basic features (e.g., difference in shape, color, brightness) and that visual information occupying the position of the highest saliency across stimulus dimensions is exogenously passed on to the "central representation" that is responsible for further stimulus analysis. Alternative explanations of the present findings and tentative speculations resulting from the present approach are discussed.

16. DESCRIPTORS	IDENTIFIERS
Reaction Time Search Vision	

17a. SECURITY CLASSIFICATION (OF REPORT)	17b. SECURITY CLASSIFICATION (OF PAGE)	17c. SECURITY CLASSIFICATION (OF ABSTRACT)
18. DISTRIBUTION/AVAILABILITY STATEMENT Unlimited availability		17d. SECURITY CLASSIFICATION (OF TITLES)

VERZENDLIJST

1. Hoofddirecteur van TNO-Defensieonderzoek

2. Directie Wetenschappelijk Onderzoek en Ontwikkeling Defensie

3. { Hoofd Wetenschappelijk Onderzoek KL
   Plv. Hoofd Wetenschappelijk Onderzoek KL

4. Hoofd Wetenschappelijk Onderzoek KLu

5. { Hoofd Wetenschappelijk Onderzoek KM
   Plv. Hoofd Wetenschappelijk Onderzoek KM

6, 7, 8. Hoofd van het Wetensch. en Techn. Doc.- en Inform. Centrum voor de Krijgsmacht

Extra exemplaren van dit rapport kunnen worden aangevraagd door tussenkomst van de HWOs of de DWOO.